On the form of witness terms

Stefan Hetzl

Archive for Mathematical Logic manuscript No.
(will be inserted by the editor)

On the form of witness terms

Stefan Hetzl

Received: date / Accepted: date

Abstract We investigate the development of terms during cut-elimination in first-order logic and Peano arithmetic for proofs of existential formulas. The form of witness terms in cut-free proofs is characterized in terms of structured combinations of basic substitutions. Based on this result, a regular tree grammar computing witness terms is given and a class of proofs is shown to have only elementary cut-elimination.

Keywords Cut-elimination · Herbrand's theorem · Peano arithmetic

Mathematics Subject Classification (2000) 03F05 · 03F07 · 03F30

1 Introduction

Cut-elimination is a tool of central importance for proof theory. It has traditionally been used to prove meta-theorems, in particular consistency-results. The situation is similar for related methods like normalization [30], Gödel's Dialectica interpretation [20] or Hilbert's ε-calculus [23]. However, these methods can also be applied to formalized mathematical proofs to extract constructive information, for example a program, from them [24,33]. Gentzen's original cut-elimination proof [15] consists essentially of a set of proof rewrite rules and a terminating strategy for applying these rules. The same is true about most cut-elimination theorems since: apart from the strategy of picking an uppermost cut for reduction as in [15], also picking a lowermost cut [16] or

This work was supported by INRIA and by a Marie Curie Intra European Fellowship within the 7th European Community Framework Programme.

Stefan Hetzl
Laboratoire Preuves, Programmes et Systèmes (PPS)
Université Paris Diderot
175 Rue du Chevaleret, 75013 Paris, France
E-mail: stefan.hetzl@pps.jussieu.fr

one of maximal logical complexity [35] turned out to be useful for obtaining termination. Several restrictions of the general proof rewrite rules with the aim of obtaining, in addition to termination, a confluence property have also been investigated [10,11]. Each restriction of the full set of proof rewrite rules has the (sometimes intended) effect of limiting the obtainable results. However, recent work [1] has shown that the number of (significantly different) normal forms may increase non-elementarily in the size of the original proof. An investigation of cut-elimination as non-deterministic computation can be found in [37,38], including a case study of a non-confluent proof in [37]. A cut-elimination method which produces even more normal forms than any method based on proof rewriting is [5], with case studies exhibiting non-confluent behavior in [2,3]. Another investigation extracting different algorithms from a classical proof is [32]. In general it is far from clear which cut-free proofs can and which cannot be obtained from a given proof with cuts. The present investigation is motivated by the interest in a characterization of the obtainable cut-free proofs.

The first aspect of this question to be dealt with is to make precise *what* is to be characterized as clearly there are aspects of formal proofs which are mathematically uninteresting. Here we will – along the lines of Herbrand's theorem – restrict our attention to the term level of a first-order proof. Let $F = \exists x_1 \cdots \exists x_n A(x_1, \ldots, x_n)$ be a valid existential formula and $\bar{t}_1, \bar{t}_2, \ldots$ be an enumeration of all n-tuples of variable-free terms in the considered language, then

$$\mathcal{H}(F) = \{\ \{A(\bar{t}_i) \mid i \in I\}\ \mid\ I \subseteq \mathbb{N},\ \bigvee_{i \in I} A(\bar{t}_i) \text{ tautology}\ \}$$

is an upper semi-lattice with the set of all instances of $A(\bar{x})$ as the unique maximal element and those sets of instances as minimal elements where removing a single formula renders them a non-tautology. Each proof π of F with cuts induces a set of points in $\mathcal{H}(F)$: the cut-free proofs reachable by cut-elimination from π. What shall be characterized then is the *least upper bound* of the reachable proofs: it represents the content of π on an elementary level (in the sense of the subformula property) but at the same time it considers π in its full generality (as no particular proof has been pre-determined by the choice of a cut-elimination procedure). The second aspect is to fix *in terms of what* this least upper bound is to be characterized: being interested in a set of term-tuples, what has to be sought is a characterization based on a pure term formalism that does not refer to proofs.

In this paper we give a characterization of a non-trivial (but not the least) upper bound in terms of a regular tree grammar [9,14]. As an application we obtain a certain class of proofs having only elementary cut-elimination. In the second part of the paper these results are extended to Peano arithmetic and demonstrated on the formalization of a short proof in number theory. From the algorithmic point of view, this method provides a new way of computing witness terms that circumvents cut-elimination.

2 Cut-elimination and Herbrand-disjunctions

Definition 1 A sequent is a pair of multisets of formulas. A proof is a tree that starts with sequents of the form $A \to A$ for an atomic formula A and is built up using the following rules.

$$\frac{\Gamma \to \Delta, A \quad \Pi \to \Lambda, B}{\Gamma, \Pi \to \Delta, \Lambda, A \wedge B}\ \wedge_{\mathrm{r}} \qquad \frac{A, \Gamma \to \Delta \quad B, \Pi \to \Lambda}{A \vee B, \Gamma, \Pi \to \Delta, \Lambda}\ \vee_{\mathrm{l}}$$

$$\frac{A, B, \Gamma \to \Delta}{A \wedge B, \Gamma \to \Delta}\ \wedge_{\mathrm{l}} \quad \frac{\Gamma \to \Delta, A, B}{\Gamma \to \Delta, A \vee B}\ \vee_{\mathrm{r}} \quad \frac{\Gamma \to \Delta, A}{\neg A, \Gamma \to \Delta}\ \neg_{\mathrm{l}} \quad \frac{A, \Gamma \to \Delta}{\Gamma \to \Delta, \neg A}\ \neg_{\mathrm{r}}$$

$$\frac{A(t), \Gamma \to \Delta}{(\forall x)A(x), \Gamma \to \Delta}\ \forall_{\mathrm{l}} \qquad \frac{\Gamma \to \Delta, A(y)}{\Gamma \to \Delta, (\forall x)A(x)}\ \forall_{\mathrm{r}}$$

$$\frac{\Gamma \to \Delta, A(t)}{\Gamma \to \Delta, (\exists x)A(x)}\ \exists_{\mathrm{r}} \qquad \frac{A(y), \Gamma \to \Delta}{(\exists x)A(x), \Gamma \to \Delta}\ \exists_{\mathrm{l}}$$

$$\frac{\Gamma \to \Delta}{A, \Gamma \to \Delta}\ \mathrm{w_l} \qquad \frac{\Gamma \to \Delta}{\Gamma \to \Delta, A}\ \mathrm{w_r} \qquad \frac{A, A, \Gamma \to \Delta}{A, \Gamma \to \Delta}\ \mathrm{c_l} \qquad \frac{\Gamma \to \Delta, A, A}{\Gamma \to \Delta, A}\ \mathrm{c_r}$$

$$\frac{\Gamma \to \Delta, A \quad A, \Pi \to \Lambda}{\Gamma, \Pi \to \Delta, \Lambda}\ \mathrm{cut}$$

The quantifier rules are subject to the usual conditions:

1. t must not contain a variable which is bound in A,
2. y is called *eigenvariable* and must not occur in $\Gamma \cup \Delta \cup \{A\}$ (eigenvariable condition).

For the sake of technical simplification, we restrict our attention to proofs of Σ_1-sentences, i.e. to sequents of the form $\to \overline{\exists x}F$ where F is quantifier-free and $\overline{\exists x}F$ contains no free variables. As the following proposition shows, this is not a severe restriction. Let $|\pi|$ denote the number of sequents in the proof π.

Proposition 1 *For any sequent s there is a Σ_1-sentence F which is valid iff s is. Furthermore, for each proof π of s there is a proof π' of F with $|\pi'| = \mathrm{O}(|\pi|^2)$.*

Proof By skolemizing the proof π we obtain a proof π_1 of a sequent s_1 which does not contain strong quantifiers and $|\pi_1| \leq |\pi|$, see [4]. The proof π' is defined by first appending to π_1 several $\neg_{\mathrm{r}}$- and $\vee_{\mathrm{r}}$-inferences to combine all formulas in s_1 into a single one and then cutting with quantifier shifts to arrive at the prenex form F. Free variables in s are treated as constants in F and the result on the formulas follows from skolemization being validity-preserving.

For complexity-reasons it is advisable to carry out the above transformation by skolemizing first and prenexifying afterwards (see [4]).

Definition 2 Let π be a proof of a Σ_1-sentence and ψ be a subproof of π. The *Herbrand-set* $\mathrm{H}(\psi, \pi)$ of ψ w.r.t. π is defined as follows. If ψ is a quantifier-free axiom $A \to A$, then

$$\mathrm{H}(\psi, \pi) := \emptyset.$$

If ψ is of the form

$$\dfrac{\begin{matrix}(\psi')\\ \Gamma \to \Delta, F(t)\end{matrix}}{\Gamma \to \Delta, \exists x F(x)}\ \exists_{\mathrm{r}}$$

where F is quantifier-free and $\exists x F(x)$ is ancestor of the formula in the end-sequent of π, then

$$\mathrm{H}(\psi, \pi) := \mathrm{H}(\psi', \pi) \cup \{F(t)\}.$$

If ψ ends with any other quantifier inference or if ψ ends with a unary inference which is not a quantifier inference, let ψ' be the immediate subproof of ψ and define

$$\mathrm{H}(\psi, \pi) := \mathrm{H}(\psi', \pi).$$

If ψ ends with a binary rule, let ψ_1 and ψ_2 be the two immediate subproofs of ψ and define

$$\mathrm{H}(\psi, \pi) := \mathrm{H}(\psi_1, \pi) \cup \mathrm{H}(\psi_2, \pi).$$

We write $\mathrm{H}(\pi)$ for $\mathrm{H}(\pi, \pi)$. For cut-free π, the formula $\bigvee \mathrm{H}(\pi)$ is a tautology, which is Gentzen's form of Herbrand's theorem, the mid-sequent theorem.

Example 1 Let P, Q, R be unary predicate symbols and define the proof $\pi =$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\to P(a), P(b)}\ \exists_{\mathrm{r}}}{\to \exists x P(x), P(b)}\ \exists_{\mathrm{r}}}{\to \exists x P(x), \exists x P(x)}\ \mathrm{c}_{\mathrm{r}}}{\to \exists x P(x)} \qquad \dfrac{\dfrac{\dfrac{}{P(\alpha) \to Q(f(\alpha))}\ \exists_{\mathrm{r}}}{P(\alpha) \to \exists x Q(x)} \qquad \dfrac{\dfrac{}{P(\alpha), Q(\beta) \to R(g(\alpha,\beta))}\ \exists_{\mathrm{r}}}{P(\alpha), Q(\beta) \to \exists x R(x)}\ \exists_{\mathrm{l}}}{P(\alpha), \exists x Q(x) \to \exists x R(x)}\ \mathrm{c}_{\mathrm{l}}, \mathrm{cut}}{\dfrac{P(\alpha) \to \exists x R(x)}{\exists x P(x) \to \exists x R(x)}\ \exists_{\mathrm{l}}}\ \mathrm{cut}$$
$$\to \exists x R(x)$$

in the sequent calculus extended by the initial sequents of π as additional axiom sequents. Then $\mathrm{H}(\pi) = \{R(g(\alpha, \beta))\}$.

It will turn out to be useful to have a mechanism that keeps track of variable names. We assume that the set of free variables is partitioned into infinitely many classes, each with infinitely many elements. Each class has exactly one distinguished element which will be called *initial variable*. For a free variable x we write $\iota(x)$ to denote the initial variable of the class x belongs to. As a convention, proofs at the beginning of cut-elimination sequences will only contain initial variables.

A quantifier occurrence in a sequent is called strong if it is positive $\forall$ or negative $\exists$ and weak otherwise. A proof is called regular if the strong quantifier rules have pairwise different eigenvariables. Cut-elimination is a set of rewrite

rules transforming regular proofs into regular proofs. If the cut formula is introduced by quantifier rules on both sides immediately above the cut, then

$$\dfrac{\dfrac{\begin{matrix}(\pi_1)\\ \Gamma \to \Delta, A(t)\end{matrix}}{\Gamma \to \Delta, \exists x A(x)}\,\exists_{\mathrm{r}} \quad \dfrac{\begin{matrix}(\pi_2)\\ A(y), \Pi \to \Lambda\end{matrix}}{\exists x A(x), \Pi \to \Lambda}\,\exists_{\mathrm{l}}}{\Gamma, \Pi \to \Delta, \Lambda}\,\mathrm{cut}$$

$\mapsto$

$$\dfrac{\begin{matrix}(\pi_1)\\ \Gamma \to \Delta, A(t)\end{matrix} \quad \begin{matrix}(\pi_2[y \leftarrow t])\\ A(t)\Pi \to \Lambda\end{matrix}}{\Gamma, \Pi \to \Delta, \Lambda}\,\mathrm{cut}$$

If the cut formula is introduced by a contraction in a proof π of the form

$$\dfrac{\dfrac{\begin{matrix}(\pi_1)\\ \Gamma \to \Delta, A, A\end{matrix}}{\Gamma \to \Delta, A}\,\mathrm{c}_{\mathrm{r}} \quad \begin{matrix}(\pi_2)\\ A, \Pi \to \Lambda\end{matrix}}{\Gamma, \Pi \to \Delta, \Lambda}\,\mathrm{cut}\ ,$$

let $\{x_1, \ldots, x_n\}$ be the eigenvariables introduced by strong quantifier rules in π_2, let $\{x'_1, x''_1, \ldots, x'_n, x''_n\}$ be fresh variables s.t. $\iota(x'_i) = \iota(x''_i) = \iota(x_i)$ for $i = 1, \ldots, n$ and define

$$\pi \quad \mapsto \quad \dfrac{\dfrac{\dfrac{\begin{matrix}(\pi_1)\\ \Gamma \to \Delta, A, A\end{matrix} \quad \begin{matrix}(\pi_2[x_i \leftarrow x'_i]_{i=1}^n)\\ A, \Pi \to \Lambda\end{matrix}}{\Gamma, \Pi \to \Delta, \Lambda, A}\,\mathrm{cut} \quad \begin{matrix}(\pi_2[x_i \leftarrow x''_i]_{i=1}^n)\\ A, \Pi \to \Lambda\end{matrix}}{\Gamma, \Pi, \Pi \to \Delta, \Lambda, \Lambda}\,\mathrm{cut}}{\Gamma, \Pi \to \Delta, \Lambda}\,\mathrm{c}^*\ .$$

As a convention, we assume that the x'_i and x''_i are chosen so as to be fresh not only for the current proof but for all proofs up to the current one in the cut-elimination sequence under consideration. In addition to the above, $\mapsto$ also contains the usual rules for permuting cuts upwards, removing propositional top-level symbols from the cut formula and for removing cuts with axioms, for the complete list of rules see Appendix A.1. With $\to$ we denote the compatible closure of $\mapsto$ and with $\twoheadrightarrow$ the reflexive and transitive closure of $\to$.

For our purposes, it will be convenient to associate a set of substitutions to a cut-elimination sequence as in [22]. If $\pi \to \pi'$ is a quantifier-reduction as above, we associate the singleton set $\{[y \leftarrow t]\}$ to this step which is written as $\pi \to^{\{[y \leftarrow t]\}} \pi'$. If $\pi \to \pi'$ is a contraction-reduction of the above form, then the reduction with associated substitution-set is $\pi \to^{\{[x_i \leftarrow x'_i]_{i=1}^n, [x_i \leftarrow x''_i]_{i=1}^n\}} \pi'$. To any other reduction the singleton set $\{\mathrm{id}\}$ is associated. To a cut-elimination sequence $\pi_1 \to^{\Sigma_1} \pi_2 \to^{\Sigma_2} \cdots \to^{\Sigma_n} \pi_{n+1}$ we associate $\Sigma := \Sigma_1 \cdots \Sigma_n$ where the concatenation of two sets of substitutions is defined as $\Sigma\Theta := \{\sigma\theta \mid \sigma \in \Sigma, \theta \in \Theta\}$ which is associative. Similarly, the application of a set of substitutions Σ to a set of formulas $\mathcal{F}$ is defined $\mathcal{F}\Sigma := \{F\sigma \mid F \in \mathcal{F}, \sigma \in \Sigma\}$. The crucial property of the substitution-set associated to a cut-elimination sequence is that it captures all changes to the first-order level of a proof in the following sense.

Proposition 2 *Let π and π^* be proofs of a Σ_1-sentence with $\pi \twoheadrightarrow^{\Sigma} \pi^*$. Then* $\mathrm{H}(\pi^*) \subseteq \mathrm{H}(\pi)\Sigma$.

Proof By induction on the length of $\pi \twoheadrightarrow^{\Sigma} \pi^*$. If $\pi = \pi^*$, then $\mathrm{H}(\pi^*) = \mathrm{H}(\pi)$ and $\Sigma = \{\mathrm{id}\}$. If $\pi \twoheadrightarrow^{\Sigma'} \pi' \rightarrow^{\Theta} \pi^*$, make a case distinction on the type of the step $\pi' \rightarrow^{\Theta} \pi^*$: if it is an axiom reduction, a rule permutation or the reduction of a propositional connective, then $\mathrm{H}(\pi^*) = \mathrm{H}(\pi')$, $\Sigma = \Sigma'$ and the result follows from the induction hypothesis. If $\pi' \rightarrow^{\Theta} \pi^*$ is the reduction of a weakening, then $\mathrm{H}(\pi^*) \subseteq \mathrm{H}(\pi')$, $\Sigma = \Sigma'$ and the result follows from the induction hypothesis.

If $\pi' \rightarrow^{\Theta} \pi^*$ is the reduction of a contraction, let $\mathrm{H}(\pi') = H_1 \cup H_2$ where H_1 contains the formulas occurring in[1] the proof π_1 (which contains the contraction) and in the context of the reduction step and H_2 those occurring in the proof π_2 (that will be duplicated). Then $\mathrm{H}(\pi^*) = H_1 \cup H_2[x_i \leftarrow x'_i]_{i=1}^n \cup H_2[x_i \leftarrow x''_i]_{i=1}^n$. On the other hand, $\mathrm{H}(\pi')\Theta = H_1[x_i \leftarrow x'_i]_{i=1}^n \cup H_2[x_i \leftarrow x'_i]_{i=1}^n \cup H_1[x_i \leftarrow x''_i]_{i=1}^n \cup H_2[x_i \leftarrow x''_i]_{i=1}^n$ but due to the regularity of the proof, the x_i do not appear in H_1 and therefore $\mathrm{H}(\pi^*) = \mathrm{H}(\pi')\Theta$ from which the result follows by induction hypothesis.

If $\pi' \rightarrow^{\Theta} \pi^*$ is the reduction of a quantifier, let $\mathrm{H}(\pi') = H_1 \cup H_2$ where H_1 contains the formulas occurring in π_1 and the context of the reduction step and H_2 those occurring in π_2. Then $\mathrm{H}(\pi^*) = H_1 \cup H_2[y \leftarrow t]$ and $\mathrm{H}(\pi')\Theta = H_1[y \leftarrow t] \cup H_2[y \leftarrow t]$ but due to regularity, y does not appear in H_1 and therefore $\mathrm{H}(\pi')\Theta = \mathrm{H}(\pi^*)$ and the result follows from the induction hypothesis.

By inspecting the proof above, we can observe that the changes to the first-order level of a proof are governed by two distinct but intertwined phenomena: duplication and instantiation. The property we rely upon in the proofs to come is the possibility of decomposing the first-order modifications into a chain of instantiations and duplications. This holds for a wide range of calculi including not only variants of the sequent calculus but also, e.g. the normalization procedure of natural deduction. The scope of the present analysis is thus quite general. This level of flexibility will also be crucial for the extension to Peano Arithmetic described in Section 5.

3 Structured Terms

In this section we introduce a formalism for the explicit description of substitutions on first-order terms which will be used as central technical tool in the analysis of cut-elimination. We will use two different operations, substitution and injection, for modeling duplication and instantiation respectively.

Definition 3 A *structured term* (sterm) is an expression built from first-order terms and the symbols $[,], \leftarrow, \cdot$ as follows:

1. If t is a term, then t is an sterm.

[1] In this proof, and the following, we will in the context of a particular cut-elimination step use the proof-, term- and formula-names of the definition of the cut-elimination step.

2. If t is a term, $n \geq 1$, $x_1, \ldots, x_n$ are distinct variables and $T_1, \ldots, T_n$ are sterms, then $t \cdot [x_1 \leftarrow T_1, \ldots, x_n \leftarrow T_n]$ is an sterm.

We often abbreviate $[x_1 \leftarrow T_1, \ldots, x_n \leftarrow T_n]$ as $[x_i \leftarrow T_i]_{i=1}^n$.

3.1 Substitution and Injection

Definition 4 Let T be an sterm. The term T°, the *evaluation* of T, is defined as follows:

1. If $T = t$ then $T^\circ := t$.
2. If $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$, then $T^\circ := t[x_i \leftarrow T_i^\circ]_{i=1}^n$

The *variables* of T are defined as follows:

1. If $T = t$, then $\mathrm{V}(T) := \mathrm{V}(t)$ where $\mathrm{V}(t)$ is the set of variables in t.
2. If $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$, then $\mathrm{V}(T) := \mathrm{V}(t) \cup \{x_1, \ldots, x_n\}$.

The *locally free variables* of T are defined as follows:

1. If $T = t$, then $\mathrm{LFV}(T) := \mathrm{V}(t)$.
2. If $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$, then $\mathrm{LFV}(T) := \mathrm{V}(t) \setminus \{x_1, \ldots, x_n\}$.

Let σ be a substitution. The sterm $T\sigma$ is defined as follows:

1. If $T = t$, then $T\sigma := t\sigma$.
2. If $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$, then $T\sigma := t(\sigma\restriction_{\mathrm{LFV}(T)}) \cdot [x_i \leftarrow T_i\sigma]_{i=1}^n$.

Let $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$ and S be sterms and let x be a variable. The sterm $T \odot [x \leftarrow S]$, the *injection of S at x into T*, is defined as follows:

1. If $x \notin \mathrm{LFV}(T)$, then $T \odot [x \leftarrow S] := t \cdot [x_i \leftarrow T_i \odot [x \leftarrow S]]_{i=1}^n$.
2. If $x \in \mathrm{LFV}(T)$, then

$$T \odot [x \leftarrow S] := t \cdot [x \leftarrow S, x_1 \leftarrow T_1 \odot [x \leftarrow S], \ldots, x_n \leftarrow T_n \odot [x \leftarrow S]].$$

Example 2 Let $T = f(x, y) \cdot [x \leftarrow a, y \leftarrow g(x)]$, then

$$T[x \leftarrow c] = f(x, y) \cdot [x \leftarrow a, y \leftarrow g(c)]$$

and

$$T \odot [x \leftarrow c] = f(x, y) \cdot [x \leftarrow a, y \leftarrow g(x) \cdot [x \leftarrow c]].$$

Note that $(T[x \leftarrow c])^\circ = (T \odot [x \leftarrow c])^\circ = f(a, g(c))$.

The evaluation $\circ$ of an sterm, the application of a substitution, and the injection are extended to sets of sterms $\mathcal{T}$ and sets of substitutions Σ by defining $\mathcal{T}^\circ := \{T^\circ \mid T \in \mathcal{T}\}$, $\mathcal{T}\Sigma := \{T\sigma \mid T \in \mathcal{T}, \sigma \in \Sigma\}$ and $\mathcal{T} \odot [x \leftarrow S] := \{T \odot [x \leftarrow S] \mid T \in \mathcal{T}\}$. For a substitution σ, the variable-range is defined as $\mathrm{vrge}(\sigma) := \{x \mid \text{there is a } y \neq x \text{ s.t. } x \in \mathrm{V}(y\sigma)\}$, the variable-range of a set of substitutions Σ is $\mathrm{vrge}(\Sigma) := \bigcup_{\sigma \in \Sigma} \mathrm{vrge}(\sigma)$.

Lemma 1 *Let T be an sterm, σ be a substitution with* $\mathrm{vrge}(\sigma) \cap \mathrm{V}(T) = \emptyset$*, and $\mathcal{T}$ be a set of sterms and Σ be a set of substitutions with* $\mathrm{vrge}(\Sigma) \cap \mathrm{V}(\mathcal{T}) = \emptyset$*. Then 1.* $(T\sigma)^\circ = T^\circ\sigma$ *and 2.* $(\mathcal{T}\Sigma)^\circ = \mathcal{T}^\circ\Sigma$*.*

Proof 1. is shown by a straightforward induction on the structure of T and 2. follows from 1.

Lemma 2 *Let T, S be sterms, $\mathcal{T}$ be a set of sterms and x be a variable. Then 1.* $(T \odot [x \leftarrow S])^\circ = T^\circ[x \leftarrow S^\circ]$ *and 2.* $(\mathcal{T} \odot [x \leftarrow S])^\circ = \mathcal{T}^\circ[x \leftarrow S^\circ]$*.*

Proof 1. is shown by induction on T and 2. follows from 1.

3.2 Normal and Regular Structured Terms

Definition 5 An sterm is called *normal* if each subexpression of the form $t \cdot [x_i \leftarrow T_i]_{i=1}^n$ satisfies $\{x_1, \ldots, x_n\} \subseteq \mathrm{V}(t)$. A term t is called *regular* if for all $x, y \in \mathrm{V}(t)$: $\iota(x) = \iota(y) \Rightarrow x = y$. An sterm is called *regular* if all terms appearing in it are regular.

Given any sterm T, one can obtain a normal T' with $T^\circ = T'^\circ$ by deleting parts of T. The function $\iota(\cdot)$ is extended from variables to terms by defining $\iota(f(t_1, \ldots, t_n)) = f(\iota(t_1), \ldots, \iota(t_n))$ and $\iota(c) = c$. The role of normality and regularity is to serve as preconditions for the following definition.

Definition 6 For a normal and regular sterm T define $\iota(T)$, the *projection of T to the initial variables*, as follows:

1. If $T = t$, then $\iota(T) := \iota(t)$.
2. If $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$, then $\iota(T) := \iota(t) \cdot [\iota(x_i) \leftarrow \iota(T_i)]_{i=1}^n$.

Note that $\iota(t) \cdot [\iota(x_i) \leftarrow \iota(T_i)]_{i=1}^n$ is a well-defined sterm because $i \neq j \Rightarrow \iota(x_i) \neq \iota(x_j)$ by normality and regularity of T. The next two lemmas demonstrate that normality and regularity are preserved by substitution and injection in a manner which is sufficient for our later purposes.

Lemma 3 *Let T and S be normal sterms, x be a variable and σ be a substitution. Then 1. $T\sigma$ is normal and 2. $T \odot [x \leftarrow S]$ is normal.*

Proof 1. is shown by induction on T observing that substitution does not change locally bound variables. 2. is also shown by induction on T by elaborating that the only changes to locally bound variables preserve normality.

Lemma 4 *Let T and S be regular sterms, $x, x_1, x_1', \ldots, x_n, x_n'$ be variables with $\iota(x_i) = \iota(x_i')$ for $i = 1, \ldots, n$. Then 1. $T[x_i \leftarrow x_i']_{i=1}^n$ is regular and 2. $T \odot [x \leftarrow S]$ is regular.*

Proof 1. is shown by induction on T based on the observation that substituting x_i by x_i' preserves regularity as $\iota(x_i) = \iota(x_i')$. For 2. it suffices to observe that each term appearing in $T \odot [x \leftarrow S]$ appears in T or in S.

3.3 Properties of the Projection to Initial Variables

Lemma 5 *Let T be a normal and regular sterm, let $x_1, x'_1, \ldots, x_n, x'_n$ be variables with $\iota(x_i) = \iota(x'_i)$ for $i = 1, \ldots, n$. Then $T[x_i \leftarrow x'_i]_{i=1}^n$ is normal and regular and $\iota(T[x_i \leftarrow x'_i]_{i=1}^n) = \iota(T)$.*

Proof Normality and regularity follow from Lemmas 3 and 4. The equality is then shown by induction on T.

For an sterm $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$ or $T = t$, we call t the *initial term* of T. A substitution is called *base substitution* if it is of the form $[x \leftarrow t]$ and $\{x\} \cup \mathrm{V}(t)$ contains only initial variables. For a set B of base substitutions and an sterm T we say that T *is over* B if for every subexpression $x \leftarrow t$ of T we have $[x \leftarrow t] \in B$.

Lemma 6 *Let B be a set of base substitutions. Let S and T be normal and regular sterms with $\iota(S)$ and $\iota(T)$ being over B and s being the initial term of S. Let x be a variable with $[\iota(x) \leftarrow \iota(s)] \in B$. Then $T \odot [x \leftarrow S]$ is normal and regular and $\iota(T \odot [x \leftarrow S])$ is over B.*

Proof Normality and regularity follow from Lemmas 3 and 4. The claim is then shown by induction on T demonstrating that all expressions of the from $y \leftarrow t$ that appear in $\iota(T \odot [x \leftarrow S])$ either appear in $\iota(T)$ or in $\iota(S)$ or are equal to $\iota(x) \leftarrow \iota(s)$.

4 Witness Terms in First-Order Logic

Having laid the necessary groundwork above, we now return to proofs in first-order logic. To each proof π we will associate a set of base substitutions, suitable combinations of which will then serve to describe the witness terms obtainable by cut-elimination.

Definition 7 Let π be a proof and Q be a quantifier occurrence in π. Define a set of terms $\mathrm{t}(Q)$ associated with Q as follows: if Q occurs in the main formula of a weakening, then $\mathrm{t}(Q) := \emptyset$. If Q is introduced by a quantifier inference from a term t or a variable x, then $\mathrm{t}(Q) := \{t\}$ or $\mathrm{t}(Q) := \{x\}$ respectively. If Q occurs in the main formula of a contraction and Q_1, Q_2 are the two corresponding quantifiers in the auxiliary formulas of the contraction, then $\mathrm{t}(Q) := \mathrm{t}(Q_1) \cup \mathrm{t}(Q_2)$. In all other cases Q has exactly one immediate ancestor Q' and $\mathrm{t}(Q) := \mathrm{t}(Q')$.

Let π be a proof, c be a cut in π. Write $\mathrm{Q}(c)$ for the set of pairs (Q, Q') of quantifier occurrences where Q is a strong occurrence in one cut-formula of c and Q' the corresponding weak occurrence in the other cut-formula. Define the set of base substitutions of c as $\mathrm{B}(c) := \bigcup_{(Q,Q') \in \mathrm{Q}(c)} \{[x \leftarrow t] \mid x \in \mathrm{t}(Q), t \in \mathrm{t}(Q')\}$. For $c_1, \ldots, c_n$ being the cuts in π define the base substitutions of π as $\mathrm{B}(\pi) := \bigcup_{i=1}^n \mathrm{B}(c_i)$.

Example 3 Letting π be the proof defined in Example 1 we have

$$\mathrm{B}(\pi) = \{[\alpha \leftarrow a], [\alpha \leftarrow b], [\beta \leftarrow f(\alpha)]\}.$$

The following auxiliary result is the analog of Proposition 2 for base substitutions.

Lemma 7 *Let $\pi \twoheadrightarrow^{\Sigma} \pi^*$ be a cut-elimination sequence. For all $[x \leftarrow t] \in \mathrm{B}(\pi^*)$ there is $[\iota(x) \leftarrow s] \in \mathrm{B}(\pi)$ s.t. $t \in s\Sigma$.*

Proof By induction on the length of $\pi \twoheadrightarrow^{\Sigma} \pi^*$. If $\pi = \pi^*$, then $\mathrm{B}(\pi^*) = \mathrm{B}(\pi)$ and $\Sigma = \{\mathrm{id}\}$. If $\pi \twoheadrightarrow^{\Sigma'} \pi' \rightarrow^{\Theta} \pi^*$ and $\Theta = \{\mathrm{id}\}$, then $\mathrm{B}(\pi^*) \subseteq \mathrm{B}(\pi')$, $\Sigma = \Sigma'$ and the result follows from the induction hypothesis.

If $\Theta = \{[x_i \leftarrow x'_i]_{i=1}^n, [x_i \leftarrow x''_i]_{i=1}^n\}$ then for all $[x \leftarrow t] \in \mathrm{B}(\pi^*)$ there is an $[x' \leftarrow t'] \in \mathrm{B}(\pi')$ with $t \in t'\Theta$ and $x \in x'\Theta$ which implies $\iota(x) = \iota(x')$. By induction hypothesis there is $[\iota(x) \leftarrow s] \in \mathrm{B}(\pi)$ s.t. $t' \in s\Sigma'$, and thus $t \in s\Sigma$.

If $\Theta = \{[y \leftarrow u]\}$ then for all $[x \leftarrow t] \in \mathrm{B}(\pi^*)$ there is an $[x' \leftarrow t'] \in \mathrm{B}(\pi')$ s.t. $t = t'[y \leftarrow u]$ and by induction hypothesis there is $[\iota(x) \leftarrow s] \in \mathrm{B}(\pi)$ s.t. $t' = s\Sigma'$. Therefore $t \in s\Sigma$.

We are now in the position to prove the main technical lemma: each substitution associated to a cut-elimination sequence starting at a proof π has the form of an sterm whose projection to the initial variables is over $\mathrm{B}(\pi)$. In order to describe substitutions by sterms we introduce a new function symbol $C(\cdot)$ which will represent the context in a proof. Later, we will replace $C(\cdot)$ by those members of the Herbrand-set to which the substitution shall be applied. For an sterm T with initial term t we say that T *is based on* s if $\iota(t) = s$.

Lemma 8 *Let $\pi \twoheadrightarrow^{\Sigma} \pi^*$ be a cut-elimination sequence, let $\{\alpha_1, \ldots, \alpha_m\}$ be the initial variables occurring in π. Then there is a set $\mathcal{T}$ of normal and regular sterms based on $C(\alpha_1, \ldots, \alpha_m)$ s.t. $\mathcal{T}^\circ = C(\alpha_1, \ldots, \alpha_m)\Sigma$ and $\iota(\mathcal{T})$ is over $\mathrm{B}(\pi)$.*

Proof We abbreviate $C(\alpha_1, \ldots, \alpha_m)$ as $C(\bar{\alpha})$. The result is shown by induction on the length of $\pi \twoheadrightarrow^{\Sigma} \pi^*$. If $\pi = \pi^*$ then $\Sigma = \{\mathrm{id}\}$ and $\mathcal{T} := \{C(\bar{\alpha})\}$. If $\pi \twoheadrightarrow^{\Sigma'} \pi' \rightarrow^{\Theta} \pi^*$ then by induction hypothesis there is a set $\mathcal{T}'$ of normal and regular sterms based on $C(\bar{\alpha})$ with $\mathcal{T}'^\circ = C(\bar{\alpha})\Sigma'$ and $\iota(\mathcal{T}')$ being over $\mathrm{B}(\pi)$. If $\Theta = \{\mathrm{id}\}$, then $\Sigma = \Sigma'$ and $\mathcal{T} := \mathcal{T}'$.

If $\Theta = \{[x_i \leftarrow x'_i]_{i=1}^n, [x_i \leftarrow x''_i]_{i=1}^n\}$, define $\mathcal{T} := \mathcal{T}'\Theta$ which is based on $C(\bar{\alpha})$. By Lemma 5, $\mathcal{T}$ is normal and regular and $\iota(\mathcal{T}) = \iota(\mathcal{T}'[x_i \leftarrow x'_i]_{i=1}^n) \cup \iota(\mathcal{T}'[x_i \leftarrow x''_i]_{i=1}^n) = \iota(\mathcal{T}')$ which is therefore over $\mathrm{B}(\pi)$. Furthermore, the x'_i and x''_i are not in $\mathrm{V}(\mathcal{T}')$ by the convention on the choice of variables at contraction-steps. Therefore we can apply Lemma 1 to conclude $\mathcal{T}^\circ = \mathcal{T}'^\circ\Theta = C(\bar{\alpha})\Sigma$ from the induction hypothesis.

If $\Theta = \{[x \leftarrow t]\}$, then, as $[x \leftarrow t] \in \mathrm{B}(\pi')$, by Lemma 7 there is a $[\iota(x) \leftarrow s] \in \mathrm{B}(\pi)$ and $\sigma \in \Sigma'$ s.t. $t = s\sigma$. By induction hypothesis there is a $S' \in \mathcal{T}'$ s.t. $S'^\circ = C(\bar{\alpha})\sigma$. S' being normal is of the form $S' = C(\bar{\alpha}) \cdot [\alpha_{i_j} \leftarrow S_j]_{j=1}^l$. Define $S := s \cdot [\alpha_{i_j} \leftarrow S_j]_{\alpha_{i_j} \in \mathrm{V}(s)}$ and observe that $S^\circ = s[\alpha_{i_j} \leftarrow {S_j}^\circ]_{\alpha_{i_j} \in \mathrm{V}(s)} = s\sigma =$

t. Define $\mathcal{T} := \mathcal{T}' \odot [x \leftarrow S]$ and observe that $\mathcal{T}^\circ = \mathcal{T}'^\circ[x \leftarrow S^\circ] = C(\bar{\alpha})\Sigma$ by Lemma 2 and the induction hypothesis. Furthermore S is normal and regular because the S_j are and the α_{i_j} appear in S, $\mathcal{T}'$ is normal and regular by induction hypothesis and $[\iota(x) \leftarrow \iota(s)] \in \mathrm{B}(\pi)$. Therefore, by Lemma 6, also $\mathcal{T}$ is normal and regular and $\iota(\mathcal{T})$ is over $\mathrm{B}(\pi)$.

In order to describe the Herbrand-disjuncts and not just the witness-terms, we will treat the logical connectives $\wedge, \vee, \neg$ as well as the predicate symbols as part of the term signature. Then, sterms can be evaluated to formulas. The main result for first-order logic is:

Theorem 1 *Let π be a proof of a Σ_1-sentence and π^* be a cut-free proof with $\pi \twoheadrightarrow \pi^*$. Then there is a set $\mathcal{S}$ of sterms over $\mathrm{B}(\pi)$ whose initial terms are elements of $\mathrm{H}(\pi)$ s.t. $\mathrm{H}(\pi^*) = \mathcal{S}^\circ$.*

Proof Let $\pi \twoheadrightarrow^{\Sigma} \pi^*$ and $F^* \in \mathrm{H}(\pi^*)$; then by Proposition 2 there is $F \in \mathrm{H}(\pi)$ and $\sigma \in \Sigma$ s.t. $F^* = F\sigma$. By Lemma 8 there is a normal and regular sterm S based on $C(\bar{\alpha})$ s.t. $S^\circ = C(\bar{\alpha})\sigma$ and $\iota(S)$ is over $\mathrm{B}(\pi)$. Define the sterm $T := F \cdot S_0$ where S_0 is S after dropping $C(\bar{\alpha})$ and all top-level substitutions which refer to initial variables that do not occur in F. Then T is over $\mathrm{B}(\pi)$ with initial term F and $T^\circ = \iota(F\sigma) = \iota(F^*)$ and as π^* is cut-free $\iota(F^*) = F^*$.

Every Herbrand-disjunct in a cut-free proof can thus be decomposed into a formula from the Herbrand-set of the original proof π and an sterm built up from base substitutions of π. This result gives the following upper bound.

Corollary 1 *Let π be a proof of a Σ_1-sentence F and let $\mathcal{S}$ be the set of sterms over $\mathrm{B}(\pi)$ with initial term from $\mathrm{H}(\pi)$. Then*

$$\{\mathrm{H}(\pi^*) \mid \pi^* \text{ cut-free}, \pi \twoheadrightarrow \pi^*\} \subseteq \mathcal{S}^\circ$$

From the algorithmic point of view, the above result shows that we can compute an Herbrand-disjunction from a proof π with cuts by successively generating sterms over $\mathrm{B}(\pi)$ with initial term from $\mathrm{H}(\pi)$ until we find a tautology. An advantage of such a procedure is that it allows complete freedom in the order of computation of witness terms and thus to find e.g. the shortest Herbrand-disjunction or one where the size of the witness of a certain quantifier is minimal.

Example 4 The upper bound on $\mathrm{H}(\pi^*)$ provided by Corollary 1 is not the least upper bound. Letting π be the proof defined in example 1 and $\mathcal{S}$ be the set of sterms having initial terms from $\mathrm{H}(\pi)$ and being over $\mathrm{B}(\pi)$, we have $\mathcal{S}^\circ = \{R(g(a,a)), R(g(a,b)), R(g(b,a)), R(g(b,b))\}$. On the other hand, for all cut-free π^* with $\pi \twoheadrightarrow \pi^*$ we obtain $\mathrm{H}(\pi^*) = \{R(g(a,a)), R(g(b,b))\}$. This example shows that cut-elimination does not generate all sterms but only such that satisfy certain structural restrictions. Describing these restrictions is left to future work.

4.1 Tree Grammars

The above characterization of witness terms uses the somewhat non-standard notion of structured term because it is well-adapted to the changes induced by cut-elimination. We will now derive from it a characterization in terms of a regular tree grammar. Regular tree languages are a natural generalization of regular (string) languages and are among the standard notions in the theory of formal languages [14]. A regular tree language can be described in several different, but equivalent ways, in particular as automaton or as grammar. We choose the presentation as grammar because the derivation of trees from a grammar closely resembles the effect of cut-elimination on the Herbrand-set. We follow the notation of [9]. For Σ being a set of symbols with associated arities, let $\mathrm{T}(\Sigma)$ denote the set of terms – or equivalently: trees – over Σ.

Definition 8 A *regular tree grammar* is a quadruple $G = (\alpha, N, F, R)$ composed of an *axiom* α, a set N of *non-terminal symbols* with $\alpha \in N$, a set F of *terminal symbols* with $F \cap N = \emptyset$ and a set R of production rules of the form $\beta \to t$ where $\beta \in N$ and $t \in \mathrm{T}(F \cup N)$.

Note that – in contrast to string grammars – a terminal symbol comes with an associated arity allowing the formation of trees. The non-terminal symbols however all have arity 0. Given a regular tree grammar $G = (\alpha, N, F, R)$, the *derivation relation* $\to_G$ associated to G is defined for $s, t \in \mathrm{T}(F \cup N)$ as $s \to_G t$ if there is a production rule $\beta \to u$ and a context $r[]$ s.t. $s = r[\beta]$ and $t = r[u]$. The *language generated by* G is $\mathrm{L}(G) := \{t \in \mathrm{T}(F) \mid \alpha \twoheadrightarrow_G t\}$ where $\twoheadrightarrow_G$ is the reflexive and transitive closure of $\to_G$. Given a proof π, let $\Sigma(\pi)$ denote the set of symbols consisting of the term-signature of π, the predicate-signature of π and the propositional connectives.

Definition 9 Let π be a proof of a Σ_1-sentence containing the initial variables $\{\alpha_1, \ldots, \alpha_n\}$. The grammar $\mathrm{G}(\pi) = (\varphi, N, F, R)$ of π is defined by setting $N = \{\varphi, \alpha_1, \ldots, \alpha_n\}$, $F = \Sigma(\pi)$ and

$$R = \{\varphi \to F \mid F \in \mathrm{H}(\pi)\} \cup \{\alpha \to t \mid [\alpha \leftarrow t] \in \mathrm{B}(\pi)\}.$$

Example 5 Letting π be the proof defined in Example 1 we have $\mathrm{G}(\pi) = (\varphi, N, F, R)$ with axiom φ, non-terminal symbols $N = \{\varphi, \alpha, \beta\}$, terminal symbols $F = \{P, Q, R, f, g, \wedge, \vee, \neg\}$ and production rules $R = \{\varphi \to R(g(\alpha, \beta)), \alpha \to a, \alpha \to b, \beta \to f(\alpha)\}$.

Lemma 9 *Let π be a proof of a Σ_1-sentence. For any sterm T over $\mathrm{B}(\pi)$ with initial term t we have $t \twoheadrightarrow_{\mathrm{G}(\pi)} T^\circ$.*

Proof If $T = t$, the result is trivially true. If $T = t \cdot [x_i \leftarrow T_i]_{i=1}^n$, then letting t_i be the initial term of T_i, we have $t_i \twoheadrightarrow_{\mathrm{G}(\pi)} T_i^\circ$ by induction hypothesis. But there are production rules $x_i \to t_i$ in $\mathrm{G}(\pi)$ and thus $x_i \twoheadrightarrow_{G(\pi)} T_i^\circ$ which when applied to all occurrences of x_i in t gives the result.

Corollary 2 *Let π be a proof of a Σ_1-sentence. Let π^* be a cut-free proof with $\pi \twoheadrightarrow \pi^*$. Then $\mathrm{H}(\pi^*) \subseteq \mathrm{L}(\mathrm{G}(\pi))$.*

Proof Let $F^* \in \mathrm{H}(\pi^*)$, then by Theorem 1 there is an sterm T over $\mathrm{B}(\pi)$ with initial term $F \in \mathrm{H}(\pi)$ s.t. $F^* = T^\circ$ and thus by Lemma 9, $F \twoheadrightarrow_{\mathrm{G}(\pi)} F^*$ which – as π^* is cut-free – implies $F^* \in \mathrm{L}(\mathrm{G}(\pi))$.

Note that Theorem 1 is slightly stronger than Corollary 2 and differs from it on terms containing several occurrences of the same variable. For example, if $H(\pi) = \{F(\alpha, \alpha)\}$ and $\mathrm{B}(\pi) = \{[\alpha \leftarrow a], [\alpha \leftarrow b]\}$, then the only sterms admitted by Theorem 1 generate $F(a, a)$ and $F(b, b)$ while the grammar generates $F(a, b)$ and $F(b, a)$ in addition.

4.2 Acyclic and Directed Proofs

It is a well-known result that the worst-case complexity of cut-elimination is non-elementary. Lower bounds have been given in [29,31,34]. In this section, we will use Theorem 1 to show that a certain class of proofs, *acyclic proofs*, has an only elementary cut-elimination. Upper bounds for the general problem based on the depth of cut-formulas can be found in [33,19,6]. W. Zhang has improved these upper bounds in [39] by using the number of nested quantifiers (nqf) instead and further in [40] by using the number of alternations between $\forall - \wedge$ and $\exists - \vee$-blocks (aqf). These results have been further improved by P. Gerhardy in [17,18] by considering, in addition to the cut-formulas, the way they are used in the proof: a part of a cut-formula which is not contracted can be eliminated with only exponential expense, regardless of the connectives that appear in this part. He introduced the measures of contracted nested quantifier depth (cnqf) and contracted alternating quantifier depth (caqf). In all of these cases, a fixed bound on the measure immediately translates to a fixed bound on the number of iterations of the exponential function and thus to elementary cut-elimination.

The following acyclicity-condition can also be viewed as extending the complexity-analysis of cut-elimination from the cut-formulas to the way they are used in the proof. Let B be a set of base substitutions and let $x, y \in \mathrm{dom}(B)$. Write $y <^1 x$ if there is a $\sigma \in B$ s.t. $y \in \mathrm{V}(x\sigma)$; write $<$ for the transitive closure of $<^1$. B is called cyclic if there is an $x \in \mathrm{dom}(B)$ s.t. $x < x$ and acyclic otherwise. A proof π is called cyclic iff $\mathrm{B}(\pi)$ is.

Corollary 3 *Let π be an acyclic proof of a Σ_1-sentence and let π^* be cut-free with $\pi \twoheadrightarrow \pi^*$. Then $|\mathrm{H}(\pi^*)| \leq |\pi|^{|\pi|^{|\pi|+1}}$*

Proof If a variable $x \in \mathrm{dom}(\mathrm{B}(\pi))$ is $<_{\mathrm{B}(\pi)}$-minimal, then define $\mathrm{rank}(x) := 1$, if x is not $<_{\mathrm{B}(\pi)}$-minimal, define $\mathrm{rank}(x) := \max\{\mathrm{rank}(y) \mid y <^1_{\mathrm{B}(\pi)} x\}+1$. Due to acyclicity, the rank of a variable is well-defined. Let r be the maximal rank of all variables of $\mathrm{dom}(\mathrm{B}(\pi))$. Let m be the maximal number of substitutions in $\mathrm{B}(\pi)$ having the same left side and let v be the number of initial variables in π. We first find an upper bound on $N_k :=$

$$\max_{x \in \mathrm{dom}(\mathrm{B}(\pi))} |\{[x \leftarrow T] \mid \mathrm{rank}(x) = k, x \cdot [x \leftarrow T] \text{ normal sterm over } \mathrm{B}(\pi)\}|$$

for $k = 1, \ldots, r$. First $N_1 \leq m$; secondly, $N_{k+1} \leq m \cdot N_k^v$ because after having chosen one of at most m possibilities for the leftmost substitution, it remains to choose, for each of at most v different variables, an sterm with rank k. We obtain $N_k \leq m^{\sum_{i=1}^{k} v^{i-1}}$. Let now $\mathcal{S}$ be the set of all normal sterms over $\mathrm{B}(\pi)$ having initial terms from $\mathrm{H}(\pi)$. Then

$$|\mathcal{S}| \leq |\mathrm{H}(\pi)| \cdot N_r^v \leq |\mathrm{H}(\pi)| \cdot m^{\sum_{i=1}^{r} v^i}.$$

Now $\mathrm{H}(\pi) \leq |\pi|$, $m \leq |\pi|$ being bound by the number of weak quantifier inferences corresponding to a certain quantifier in a cut, and $r \leq v \leq |\pi|$ and thus

$$|\mathcal{S}| \leq |\pi| \cdot |\pi|^{\sum_{i=1}^{|\pi|} |\pi|^i} \leq |\pi|^{|\pi|^{|\pi|+1}}.$$

Finally, by Theorem 1, $|\mathrm{H}(\pi^*)| \leq |\mathcal{S}|$.

The above result improves any upper bound based on the logical structure of cut-formulas, in particular nqf and aqf, but also cnqf and caqf as follows: Let π_n be any worst-case sequence and define π'_n from π_n by removing the term-level thus rendering every predicate nullary and every quantifier vacuous. The logical structure of the cut-formulas (as well as the contractions in the proof) do not change but the set of base substitutions becomes empty hence acyclic and thus π'_n is recognized as having elementary cut-elimination.

In order to obtain a more meaningful comparison with the known upper bounds we will now consider a class of formulas whose use for cuts induces only acyclic proofs. A cut is called directed if its cut formula does not contain both strong and weak quantifiers. A proof is called directed if all its cuts are.

Lemma 10 *Every directed proof is acyclic.*

Proof By induction on the number of cuts in the proof. All cut-free proofs are acyclic. For the induction step, consider a proof π and let ι be the lowest binary inference with subproofs π_1 and π_2 s.t. either 1. ι is a cut or 2. both π_1 and π_2 contain a cut. In case 2, $<_\pi = <_{\pi_1} \cup <_{\pi_2}$ which is acyclic by induction hypothesis. In case 1, $<^1_\pi = <^1_{\pi_1} \cup <^1_{\pi_2} \cup <^1_{\mathrm{B}(\iota)}$. By induction hypothesis, $<_{\pi_1}$ and $<_{\pi_2}$ are acyclic and as ι is directed, also $<_{\mathrm{B}(\iota)}$ is acyclic. Therefore, a cycle in $<_\pi$ must be of the form $x_1 \leq_{\pi_1} x_2 <_{\mathrm{B}(\iota)} y_1 \leq_{\pi_2} y_2 <_{\mathrm{B}(\iota)} x_1$ where x_1, x_2 are eigenvariables of strong quantifier inferences in π_1 and y_1, y_2 of inferences in π_2. However, as ι is directed, only one of $x_2 <_{\mathrm{B}(\iota)} y_1$ and $y_2 <_{\mathrm{B}(\iota)} x_1$ is possible.

Therefore the elementary upper bound of Corollary 3 applies to directed proofs. For the sake of comparison we restrict our attention to formulas in negation normal form and find for A with $\mathrm{aqf}(A) = 0$ that A is directed. For each $k > 0$ one can find directed formulas with aqf $= k$ (by alternating $\forall$ and $\vee$, or $\exists$ and $\wedge$ respectively) as well as undirected formulas (by alternating $\forall$ and $\exists$). The measure nqf behaves analogously. The bound on directed proofs thus improves the known upper bounds by exhibiting an additional class of formulas that has only elementary cut-elimination. In how far the restrictions

on the use of contractions considered in [17,18] imply properties of the graph $(\mathrm{B}(\pi), <^1)$ and vice versa is an interesting question left open for future work.

Corollary 3 shows that, in particular, the worst case sequences of [29,31, 34] with cuts are cyclic. It is also interesting to compare this result with the one obtained by A. Carbone in [8]: the logical flow graph of a short proof of the feasibility of a large number must necessarily be cyclic. This property of logical flow graphs is not very robust w.r.t. changes in the calculus. Indeed, in [7] a sequent calculus ALK (acyclic LK) is defined in which all logical flow graphs are acyclic while ordinary LK-proofs can be translated to ALK-proofs with only elementary increase in length. The cyclicity-property of the base substitutions is considerably more robust to changes of the calculus, in particular the results of this paper also hold in ALK.

The base substitutions can be regarded as a flow graph-like structure if we consider the graph whose vertices are the quantifier occurrences in a proof, and we draw a (cut-)link between two quantifier occurrences if they are ancestors of the dual occurrences in the same cut (which corresponds to the definition of $\mathrm{B}(c)$ from $\mathrm{Q}(c)$) and an (axiom-)link between Q and Q' if the term of the weak quantifier Q' contains the eigenvariable of the strong quantifier Q. Such a presentation in the framework of proof-nets has been given in [21] and [27].

Another interesting aspect of the situation is that, even though an acyclic proof has only short Herbrand-disjunctions, it may nevertheless have normal forms of arbitrary size (this can easily be seen by incorporating the double-contraction example found e.g. in [11,13,37] and in a similar form in [41] into an acyclic proof). The large normal forms of acyclic proofs are therefore only due to repetitions of the same formulas and thus mathematically meaningless. The analogous question for cyclic proofs is open: it has been shown in [1] that a (cyclic) proof can have a non-elementary number of reachable Herbrand-disjunctions. It is unclear however whether there exists a (cyclic) proof having infinitely many reachable Herbrand-disjunctions.

5 Extension to Peano Arithmetic

In this section, the above results are extended to Peano arithmetic. The language contains a symbol for every primitive recursive function. The calculus is extended by their defining equations as additional axiom sequents. Terms of the form $0, 0', 0'', \ldots$ are called numerals; we use the notation $\bar{n}$ for the numeral denoting the natural number n. We further add the induction rule

$$\frac{\Gamma \to \Delta, F(0) \quad F(y), \Pi \to \Lambda, F(y')}{\Gamma, \Pi \to \Delta, \Lambda, F(t)}\ \mathrm{ind}$$

where F is an arbitrary formula, t is an arbitrary term and y is an eigenvariable, i.e. it is not allowed to occur in $\Gamma, \Pi \to \Delta, \Lambda, F(t)$. For a variable-free term t, let $|t|$ denote its value in $\mathbb{N}$. We also add the evaluation rules

$$\frac{\Gamma \to \Delta, F(s)}{\Gamma \to \Delta, F(t)}\ \mathrm{v_r} \qquad \text{and} \qquad \frac{F(s), \Gamma \to \Delta}{F(t), \Gamma \to \Delta}\ \mathrm{v_l}$$

for variable-free terms s and t with $|s| = |t|$. Note that these rules are redundant w.r.t. provability, i.e. $F(s) \to F(t)$ is also provable without them. However, these rules (as well as the particular form of the induction rule above that differs from the one in [36]) will permit some technical simplifications later on. The proof reduction steps for cut-elimination are extended by the following rules for eliminating inductions: If t is a variable-free term with $|t| = 0$, then

$$\dfrac{\overset{(\pi_1)}{\Gamma \to \Delta, F(0)} \quad \overset{(\pi_2)}{F(y), \Pi \to \Lambda, F(y')}}{\Gamma, \Pi \to \Delta, \Lambda, F(t)}\ \mathrm{ind} \qquad \mapsto \qquad \dfrac{\dfrac{\overset{(\pi_1)}{\Gamma \to \Delta, F(0)}}{\Gamma \to \Delta, F(t)}\ \mathrm{v_r}}{\Gamma, \Pi \to \Delta, \Lambda, F(t)}\ \mathrm{w}^*.$$

If $|t| = n + 1$, then let $x_1, \ldots, x_m$ be the eigenvariables of π_2, let $x_j^{(i)}$ for $j = 1, \ldots, m$ and $i = 1, \ldots, n$ be fresh variables s.t. $\iota(x_j^{(i)}) = \iota(x_j)$. Define $\theta_i := [y \leftarrow \overline{i}, x_1 \leftarrow x_1^{(i)}, \ldots, x_m \leftarrow x_m^{(i)}]$ for $i = 0, \ldots, n$; then the same proof maps to

$$\dfrac{\dfrac{\overset{(\pi_1)}{\Gamma \to \Delta, F(0)} \quad \overset{(\pi_2\theta_0)}{F(0), \Pi \to \Lambda, F(0')}}{\Gamma, \Pi \to \Delta, \Lambda, F(0')}\ \mathrm{cut} \quad \overset{(\pi_2\theta_1)}{F(0'), \Pi \to \Lambda, F(0'')}}{\dfrac{\dfrac{\Gamma, \Pi, \Pi \to \Delta, \Lambda, \Lambda, F(0'')}{\Gamma, \Pi \to \Delta, \Lambda, F(0'')}\ \mathrm{c}^*}{\vdots}}\ \mathrm{cut}$$
$$\dfrac{\Gamma, \Pi \to \Delta, \Lambda, F(\overline{n+1})}{\Gamma, \Pi \to \Delta, \Lambda, F(t)}\ \mathrm{v_r}$$

where the same freshness convention as in first-order logic is assumed for the $x_j^{(i)}$. To the former reduction, we associate the singleton set of substitutions $\{\mathrm{id}\}$ and to the latter we associate $\{\theta_0, \ldots, \theta_n\}$. In addition, we add proof rewrite steps for shifting evaluation rules upwards, see Appendix A.2, to which we associate the substitution set $\{\mathrm{id}\}$. We restrict our attention to proofs of sequents of the form $\to F$ where F is a Σ_1^0-sentence, i.e. a formula $\exists x\ G(x)$ where G is quantifier-free and contains no variable except x. We do not impose any restriction on the formulas used in inductions and cuts.

Definition 10 The set $\mathrm{t}(Q)$ of terms associated to a quantifier occurrence Q in a PA-proof π is defined as in Definition 7 with the following addition. If Q is in the main occurrence $F(t)$ of an induction inference, let Q_1 be the corresponding quantifier in $F(0)$ and Q_2 that in $F(y')$ and define $\mathrm{t}(Q) := \mathrm{t}(Q_1) \cup \mathrm{t}(Q_2)$.

For the base substitutions we extend Definition 7 as follows. For an induction inference d write $\mathrm{Q}(d)$ for the set of pairs (Q, Q') of quantifier occurrences s.t. Q is a strong occurrence, Q' is a weak occurrence and one of the following is true.

1. Q is in $F(0)$ and Q' is the corresponding occurrence in $F(y)$

2. Q is in $F(y)$ and Q' is the corresponding occurrence in $F(0)$
3. Q is in $F(y)$ and Q' is the corresponding occurrence in $F(y')$
4. Q is in $F(y')$ and Q' is the corresponding occurrence in $F(y)$

Define

$$\mathrm{B}(d) := \{[y \leftarrow 0], [y \leftarrow y']\} \cup \bigcup_{(Q,Q') \in \mathrm{Q}(d)} \{[x \leftarrow t] \mid x \in \mathrm{t}(Q), t \in \mathrm{t}(Q')\}$$

Let $c_1, \ldots, c_n$ be the cuts and $d_1, \ldots, d_m$ be the inductions in π and define

$$\mathrm{B}(\pi) := \bigcup_{i=1}^{n} \mathrm{B}(c_i) \cup \bigcup_{i=1}^{m} \mathrm{B}(d_i)$$

In arithmetical proofs, one is usually interested in the witness terms per se (instead of, as in the case of first-order logic, in the whole Herbrand-disjunction). We will therefore only follow the development of the witness terms during cut-elimination. For a proof π of a Σ^0_1-sentence, let Q be the existential quantifier in the end-sequent and define $\mathrm{W}(\pi) := \mathrm{t}(Q)$.

Example 6 Define $\pi =$

$$\dfrac{\dfrac{\dfrac{}{\to 2^0 = \bar{1}}\ \exists_{\mathrm{r}}}{\to \exists y\ 2^0 = y}\ \exists_{\mathrm{r}} \quad \dfrac{\dfrac{\dfrac{(\psi)}{2^\beta = \gamma \to 2^{\beta'} = \bar{2} \cdot \gamma}\ \exists_{\mathrm{r}}}{2^\beta = \gamma \to \exists y\ 2^{\beta'} = y}\ \exists_{\mathrm{l}}}{\exists y\ 2^\beta = y \to \exists y\ 2^{\beta'} = y}\ \mathrm{ind}}{\dfrac{\to \exists y\ 2^\alpha = y}{\to \forall x \exists y\ 2^x = y}\ \forall_{\mathrm{r}}}$$

where ψ is cut- and induction-free. Let $\pi(\alpha)$ denote π without its last rule. Then $\mathrm{W}(\pi(\alpha)) = \{\bar{1}, \bar{2} \cdot \gamma\}$ and $\mathrm{B}(\pi(\alpha)) = \{[\gamma \leftarrow \bar{1}], [\gamma \leftarrow \bar{2} \cdot \gamma], [\beta \leftarrow 0], [\beta \leftarrow \beta']\}$.

What follows now is an extension of the results that have been obtained for first-order logic to Peano arithmetic.

Proposition 3 *Let π be a PA-proof of a Σ^0_1-sentence. If $\pi \twoheadrightarrow^{\Sigma} \pi^*$, then $\mathrm{W}(\pi^*) \subseteq \mathrm{W}(\pi)\Sigma$.*

Proof The induction base and those cases of the induction step referring to reduction rules of first-order logic are treated as in the proof of Proposition 2. It only remains to extend the induction step to the proof rewrite steps specific to PA. So let $\pi \twoheadrightarrow^{\Sigma'} \pi' \to^{\Theta} \pi^*$, which by induction hypothesis gives $\mathrm{W}(\pi') \subseteq \mathrm{W}(\pi)\Sigma'$. If $\pi' \to^{\Theta} \pi^*$ is the permutation of an evaluation, for example

$$\dfrac{\dfrac{\Gamma \to \Delta, F(r,s)}{\Gamma \to \Delta, \exists x\ F(x,s)}\ \exists_{\mathrm{r}}}{\Gamma \to \Delta, \exists x\ F(x,t)}\ \mathrm{v}_{\mathrm{r}} \quad \to \quad \dfrac{\dfrac{\Gamma \to \Delta, F(r,s)}{\Gamma \to \Delta, F(r,t)}\ \mathrm{v}_{\mathrm{r}}}{\Gamma \to \Delta, \exists x\ F(x,t)}\ \exists_{\mathrm{r}},$$

let Q' be the existential quantifier in the left proof and Q^* the one in the right proof, observe that $\mathrm{t}(Q^*) = \mathrm{t}(Q') = \{r\}$ and thus $\mathrm{W}(\pi^*) = \mathrm{W}(\pi')$ from which the result follows by induction hypothesis.

If $\pi' \to^{\Theta} \pi^*$ is an induction-elimination, let $\mathrm{W}(\pi') = W \cup W_1 \cup W_2$ where W contains the witnesses introduced by quantifier rules in the context of the reduction step, W_1 and W_2 those introduced in the proofs π_1 and π_2 above the induction inference. If $|t| = 0$, then $\mathrm{W}(\pi^*) = W \cup W_1$ which is a subset of $\mathrm{W}(\pi)\Sigma$ by induction hypothesis. If $|t| = n+1$, then $\Theta = \{\theta_1, \ldots, \theta_n\}$ where $\theta_i = [y \leftarrow \bar{i}, x_1 \leftarrow x_1^{(i)}, \ldots, x_m \leftarrow x_m^{(i)}]$ and $\mathrm{W}(\pi^*) = W \cup W_1 \cup \bigcup_{i=0}^{n} W_2\theta_i$. On the other hand, $\mathrm{W}(\pi')\Theta = \bigcup_{i=0}^{n}(W\theta_i \cup W_1\theta_i \cup W_2\theta_i) = W \cup W_1 \cup \bigcup_{i=0}^{n} W_2\theta_i$ as, due to regularity, y and the x_i do not appear in W_1. Therefore $\mathrm{W}(\pi^*) = \mathrm{W}(\pi')\Theta \subseteq \mathrm{W}(\pi)\Sigma$ by the induction hypothesis.

Lemma 11 *Let π be a PA-proof and $\pi \twoheadrightarrow^{\Sigma} \pi^*$. For all $[x \leftarrow t] \in \mathrm{B}(\pi^*)$ there is $[\iota(x) \leftarrow s] \in \mathrm{B}(\pi)$ s.t. $t \in s\Sigma$.*

Proof As in the proof of Lemma 7 we proceed by induction on the length of $\pi \twoheadrightarrow \pi^*$ and it only remains to treat the PA-specific reductions in the induction step. So assume $\pi \to^{\Sigma'} \twoheadrightarrow^{\Theta} \pi^*$. If $\pi' \to \pi^*$ is a permutation of an evaluation inference, then $\mathrm{B}(\pi^*) = \mathrm{B}(\pi')$ as the terms $\mathrm{t}(Q)$ associated to a quantifier Q do not change. As $\Theta = \{\mathrm{id}\}$, the result follows immediately by the induction hypothesis.

If $\pi' \twoheadrightarrow^{\Theta} \pi^*$ is an induction-elimination with $|t| = 0$, then $\mathrm{B}(\pi^*) \subseteq \mathrm{B}(\pi')$ and as $\Sigma = \Sigma'$ the result follows immediately by the induction hypothesis, so let $|t| = n+1$. To show the claim, it suffices to show for all $[x \leftarrow t] \in \mathrm{B}(\pi^*)$ that

$$(*) \text{ there is a } [x' \leftarrow t'] \in \mathrm{B}(\pi') \text{ s.t. } \iota(x) = \iota(x') \text{ and } t \in t'\Theta,$$

for then, by induction hypothesis, there is a $[\iota(x) \leftarrow s] \in \mathrm{B}(\pi)$ with $t' \in s\Sigma'$ which implies that $t \in s\Sigma'\Theta = s\Sigma$. To prove (*), let d be the eliminated induction in π' and observe that

$$\mathrm{B}(\pi') = B \cup B_1 \cup B_2 \cup \mathrm{B}(d)$$

where B are the base substitutions associated to cuts and inductions in the context of the reduction step and B_1 and B_2 are those of cuts and inductions in π_1 and π_2 respectively. Furthermore, let $\Theta = \{\theta_0, \ldots, \theta_n\}$ where $\theta_i = [y \leftarrow \bar{i}, x_1 \leftarrow x_1^{(i)}, \ldots, x_m \leftarrow x_m^{(i)}]$ and let $c_0, \ldots, c_n$ be the cuts replacing the induction in π^*, then

$$\mathrm{B}(\pi^*) = B^* \cup B_1 \cup \bigcup_{i=0}^{n}\{[x\theta_i \leftarrow t\theta_i] \mid [x \leftarrow t] \in B_2\} \cup \bigcup_{i=0}^{n} B(c_i)$$

where B^* are the base substitutions associated to cuts and inductions in the context of the reduction step in π^*. We will now prove (*) for each of the above subsets of $\mathrm{B}(\pi^*)$.

If $[x \leftarrow t] \in B_1$ or $[x \leftarrow t] \in \bigcup_{i=0}^{n}\{[x\theta_i \leftarrow t\theta_i] \mid [x \leftarrow t] \in B_2\}$, then (*) is immediate from observing that $\iota(x) = \iota(x\theta_i)$. Write Q' for a quantifier in the end-sequent of the reduction step in π' and Q^* for the corresponding quantifier in the end-sequent of the reduction step in π^*. If Q', Q^* are in Γ or Δ, then $\mathrm{t}(Q^*) = \mathrm{t}(Q')$. If they are in Π or Λ, then $\mathrm{t}(Q^*) = \bigcup_{i=0}^{n} t(Q')\theta_i$. If Q', Q^* are in $F(t)$, let $\mathrm{t}(Q') = T_1 \cup T_2$ where T_1 are the terms from π_1 and T_2 those from π_2 and observe that $\mathrm{t}(Q^*) = T_2\theta_n$. Putting these cases together, we obtain that for all $t \in \mathrm{t}(Q^*)$ there is a $t' \in \mathrm{t}(Q')$ s.t. $t \in t'\Theta$ and therefore that (*) is true for all $[x \leftarrow t] \in B^*$.

If $[x \leftarrow t] \in \mathrm{B}(c_0)$, then this substitution is induced by a pair (Q, Q') of quantifier occurrences where Q is strong and Q' is weak. If Q is in π_1, then Q' is in $\pi_2\theta_0$ and $t = t'\theta_0$ where $[x \leftarrow t'] \in \mathrm{B}(\iota)$. If Q is in $\pi_2\theta_0$, then Q' is in π_1, $[x \leftarrow t] \in \mathrm{B}(\iota)$ and $t\theta_i = t$ for all i by regularity. Let now $[x \leftarrow t] \in \mathrm{B}(c_i)$ for some $i \in \{1, \ldots, n\}$ with quantifier-pair (Q, Q') with Q strong and Q' weak. If Q is in $\pi_2\theta_{i-1}$ and Q' in $\pi_2\theta_i$ then there is a $[x' \leftarrow t'] \in \mathrm{B}(\iota)$ with $t'\theta_i = t$ and $x'\theta_{i-1} = x$, i.e. $\iota(x) = \iota(x')$. If, on the other hand, Q is in $\pi_2\theta_i$ and Q' in $\pi_2\theta_{i-1}$ then there is $[x' \leftarrow t'] \in \mathrm{B}(\iota)$ with $t'\theta_{i-1} = t$ and $x'\theta_i = x$, i.e. $\iota(x') = \iota(x)$. This concludes the proof of (*) for all $[x \leftarrow t] \in \mathrm{B}(\pi^*)$ and thus the proof of the lemma.

Lemma 12 *Let π be a PA-proof and $\pi \twoheadrightarrow^{\Sigma} \pi^*$, let $\{\alpha_1, \ldots, \alpha_m\}$ be the initial variables occurring in π. Then there is a set $\mathcal{T}$ of normal and regular sterms based on $C(\bar{\alpha})$ s.t. $\mathcal{T}^\circ = C(\bar{\alpha})\Sigma$ and $\iota(\mathcal{T})$ is over* $\mathrm{B}(\pi)$.

Proof By induction on the length of $\pi \twoheadrightarrow^{\Sigma} \pi^*$. The induction base and the cases of the induction step pertaining to pure first-order logic are analogous to the proof of Lemma 8, where it is important to note that the case of quantifier reduction relies on the extension of Lemma 7 to PA in form of the above Lemma 11. It remains to treat the PA-specific cases. Let $\pi \twoheadrightarrow^{\Sigma'} \pi' \rightarrow^{\Theta} \pi^*$, then by the induction hypothesis there is a set $\mathcal{T}'$ of normal and regular sterms based on $C(\bar{\alpha})$ with $\mathcal{T}'^\circ = C(\bar{\alpha})\Sigma'$ and $\iota(\mathcal{T}')$ over $\mathrm{B}(\pi)$. If $\pi \rightarrow^{\Theta} \pi^*$ is a permutation of an evaluation or an induction-elimination $|t| = 0$, let $\mathcal{T} := \mathcal{T}'$ and observe that the result follows directly from $\Sigma = \Sigma'$. Let now $\pi' \rightarrow^{\Theta} \pi^*$ be an induction-elimination with $|t| = n + 1$ where $\Theta = \{\theta_0, \ldots, \theta_n\}$ with $\theta_i = [y \leftarrow \bar{i}, x_1 \leftarrow x_1^{(i)}, \ldots, x_m \leftarrow x_m^{(i)}]$. Define a sequence of sterms as $S_0 := 0$ and $S_{k+1} := y' \cdot [y \leftarrow S_k]$ and define

$$\mathcal{T} := \bigcup_{i=0}^{n}\{T[x_j \leftarrow x_j^{(i)}]_{j=1}^{m} \odot [y \leftarrow S_i] \mid T \in \mathcal{T}'\}.$$

As $\mathcal{T}'$ is based on $C(\bar{\alpha})$, so is $\mathcal{T}$. By Lemma 5, the $T[x_j \leftarrow x_j^{(i)}]_{j=1}^{m}$ are normal and regular and $\iota(T[x_j \leftarrow x_j^{(i)}]_{j=1}^{m}) = \iota(T)$ which is therefore over $\mathrm{B}(\pi)$. Furthermore, the S_i are normal and regular, $[\iota(y) \leftarrow 0]$ as well as $[\iota(y) \leftarrow \iota(y')]$ are in $\mathrm{B}(\pi)$, therefore the $\iota(S_i)$ are over $\mathrm{B}(\pi)$ and we can apply Lemma 6 to conclude that $\mathcal{T}$ is normal and regular and $\iota(\mathcal{T})$ is over $\mathrm{B}(\pi)$. Finally, by the

convention on the choice of fresh variables at elimination of an induction, the $x_j^{(i)}$ do not appear in $\mathcal{T}'$ and we can apply Lemma 1 to obtain

$$\begin{aligned}\mathcal{T}^\circ &= \bigcup_{i=0}^{n}\{T^\circ[x_j \leftarrow x_j^{(i)}]_{j=1}^m[y \leftarrow s^i(0)] \mid T \in \mathcal{T}'\}\\ &= \mathcal{T}'^\circ\Theta = C(\alpha)\Sigma.\end{aligned}$$

Observe that while the substitution-set associated to an induction-elimination-step is a mixture of duplication and instantiation, it can be written as concatenation of these two components. This is crucial for the above result to extend to PA without having to extend the framework on the level of structured terms. The main result and its corollaries can then be proved as in the case of first-order logic.

Theorem 2 *Let π be a PA-proof of a Σ_1^0-sentence and π^* be a cut- and induction-free proof with $\pi \twoheadrightarrow \pi^*$. Then there is a set $\mathcal{S}$ of sterms over $\mathrm{B}(\pi)$ having initial terms from $\mathrm{W}(\pi)$ s.t. $\mathrm{W}(\pi^*) = \mathcal{S}^\circ$.*

The grammar of a PA-proof is defined as in the case of first-order logic with the only difference of replacing $\mathrm{H}(\pi)$ by $\mathrm{W}(\pi)$.

Corollary 4 *Let π be a PA-proof of a Σ_1^0-sentence and π^* be a cut- and induction-free proof with $\pi \twoheadrightarrow \pi^*$. Then $\mathrm{W}(\pi^*) \subseteq \mathrm{L}(\mathrm{G}(\pi))$.*

These results can be applied to a proof π of a Π_2^0-sentence as follows: assume w.l.o.g. that π ends with a $\forall_\mathrm{r}$-inference with eigenvariable α and denote with $\pi(\alpha)$ the proof π without its last rule. We can obtain $\mathrm{B}(\pi(\alpha))$ and $\mathrm{W}(\pi(\alpha))$ just as for a proof of a Σ_1^0-sentence by regarding α as a constant symbol. For $n \in \mathbb{N}$, $\mathrm{B}(\pi(\bar{n}))$ and $\mathrm{W}(\pi(\bar{n}))$ are uniform in the sense that $\mathrm{B}(\pi(\bar{n})) = \mathrm{B}(\pi(\alpha))[\alpha \leftarrow \bar{n}]$ and $\mathrm{W}(\pi(\bar{n})) = \mathrm{W}(\pi(\alpha))[\alpha \leftarrow \bar{n}]$ and therefore also the grammars and the languages induced by them are uniform in n, i.e. $\mathrm{G}(\pi(\bar{n})) = \mathrm{G}(\pi(\alpha))[\alpha \leftarrow \bar{n}]$ and $\mathrm{L}(\mathrm{G}(\pi(\bar{n}))) = \mathrm{L}(\mathrm{G}(\pi(\alpha)))[\alpha \leftarrow \bar{n}]$.

The corollary about acyclic proofs of Σ_1-sentences also holds in PA, however the presence of an induction makes a proof cyclic. It would be possible to broaden the scope of the corollary by deleting base substitutions (and thus: their cycles) that are not reachable from the initial witness terms but we do not go into more detail here.

6 Application to a Concrete Proof

In this section, we apply the above techniques to a concrete example proof of a Π_2^0-statement in number theory: We will prove that for $m \geq 2$ and $n \geq 1$ there is a number between n and $m^2 \cdot n$ which can be written as a sum of two squares. Let $S(x)$ be a quantifier-free formula s.t. $S(\bar{n})$ is true iff there are $n_1, n_2 \in \mathbb{N}$ with $n_1^2 + n_2^2 = n$. Define $A(m,n,k) \;:=\; n < k \wedge k \leq m^2 \cdot n \wedge S(k)$.

To simplify the exposition, we formalize the hypotheses $m \geq 2$ and $n \geq 1$ by proving $\forall m \forall n \exists k\, A(m'', n', k)$. Let $\pi :=$

$$\dfrac{\dfrac{\dfrac{\vdots}{\to \bar{1} < \bar{2} \wedge \bar{2} \leq (\mu'')^2 \cdot \bar{1} \wedge S(\bar{2})}}{\to \exists k\, A(\mu'', \bar{1}, k)}\ \exists_{\mathrm{r}} \qquad \dfrac{(\mathrm{IS})}{\exists k\, A(\mu'', \nu_0', k) \to \exists k\, A(\mu'', \nu_0'', k)}\ \exists_{\mathrm{l}}}{\dfrac{\to \exists k\, A(\mu'', \nu', k)}{\to \forall m \forall n \exists k\, A(m'', n', k)}\ \forall_{\mathrm{r}}, \forall_{\mathrm{r}}}\ \mathrm{ind}$$

where $\mathrm{IS} :=$

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\nu_0'' < \kappa, A(\mu'', \nu_0', \kappa) \to A(\mu'', \nu_0'', \kappa)}}{\nu_0'' < \kappa, A(\mu'', \nu_0', \kappa) \to \exists k\, A(\mu'', \nu_0'', k)}\ \exists_{\mathrm{r}}}{A(\mu'', \nu_0', \kappa) \to \exists k\, A(\mu'', \nu_0'', k), \neg \nu_0'' < \kappa}\ \neg_{\mathrm{r}} \qquad (\mathrm{IS}')}{A(\mu'', \nu_0', \kappa) \to \exists k\, A(\mu'', \nu_0'', k)}\ \mathrm{cut}$$

and $\mathrm{IS}' :=$

$$\dfrac{\dfrac{(\mathrm{IS}_1') \qquad (\mathrm{IS}_2') \qquad (\mathrm{IS}_3')}{\neg \nu_0'' < \kappa, A(\mu'', \nu_0', \kappa) \to A(\mu'', \nu_0'', (\mu'')^2 \cdot \kappa)}\ \wedge_{\mathrm{l}}^*, \mathrm{w}_{\mathrm{l}}, \wedge_{\mathrm{r}}^*}{\neg \nu_0'' < \kappa, A(\mu'', \nu_0', \kappa) \to \exists k\, A(\mu'', \nu_0'', k)}\ \exists_{\mathrm{r}}$$

and

$$\begin{aligned} &\mathrm{IS}_1' \ \text{ proves } \ \nu_0' < \kappa \to \nu_0'' < (\mu'')^2 \cdot \kappa, \\ &\mathrm{IS}_2' \ \text{ proves } \ \neg \nu_0'' < \kappa \to (\mu'')^2 \cdot \kappa \leq (\mu'')^2 \cdot \nu_0'', \ \text{ and} \\ &\mathrm{IS}_3' \ \text{ proves } \ S(\kappa) \to S((\mu'')^2 \cdot \kappa). \end{aligned}$$

For proving $S(\kappa) \to S((\mu'')^2 \cdot \kappa)$ we may choose to rely on the characterization of the sums of two squares due to Fermat and first proved by Euler [12]: $n \in \mathbb{N}$ is a sum of two squares iff all primes $p \equiv 3 \pmod 4$ have even exponent in the factorization of n. We may also choose to prove this by a direct calculation showing that products of sums of two squares are sums of two squares. It is important to note at this point that we do not have to explicitly formalize its proof nor that of any other of the as of now unproved statements to carry out the following analysis.

Let $\pi(\mu, \nu)$ denote the above proof π without its last two $\forall_{\mathrm{r}}$-rules. The witness terms for $\exists k$ are $\mathrm{W}(\pi(\mu, \nu)) = \{\bar{2}, \kappa, (\mu'')^2 \cdot \kappa\}$ and letting ι be the displayed induction, $\mathrm{B}(\iota) = \{[\kappa \leftarrow \bar{2}], [\kappa \leftarrow \kappa], [\kappa \leftarrow (\mu'')^2 \cdot \kappa], [\nu_0 \leftarrow 0], [\nu_0 \leftarrow \nu_0']\}$ and $\mathrm{B}(\pi(\mu, \nu)) = \mathrm{B}(\iota) \cup B^*$ where B^* contains the base substitutions of all cuts and inductions in those parts of the proof that have not been formalized. We obtain a grammar $\mathrm{G}(\pi(\mu, \nu)) = (\tau, N, F, R)$ with $\tau, \kappa, \nu_0 \in N$ and R consisting of the rules

$$\begin{array}{lll} \tau \to \bar{2} & \kappa \to \bar{2} & \nu_0 \to 0 \\ \tau \to \kappa & \kappa \to \kappa & \nu_0 \to \nu_0' \\ \tau \to (\mu'')^2 \cdot \kappa & \kappa \to (\mu'')^2 \cdot \kappa & \end{array}$$

plus those induced by B^*. By applying standard pruning techniques we can simplify the above grammar: the only non-terminal symbols reachable from the axiom τ are τ and κ, thus we can delete all rules whose left-hand side is different from τ and κ including all those induced by B^*. The corresponding proof parts are computationally irrelevant and only serve the purpose of verification. Secondly, observe that τ and κ have the same right-hand sides, so we can identify them. Finally, we can delete the unproductive loop rule $\tau \to \tau$ to obtain the grammar G_μ which is equivalent to $\mathrm{G}(\pi(\mu,\nu))$, has τ as axiom and only non-terminal symbol and

$$\tau \to \bar{2} \qquad\qquad \tau \to (\mu'')^2 \cdot \tau$$

as the only production rules.

Let now $m \geq 2$ and $n \geq 1$. By inspecting G_μ we can see that eliminating the cuts in $\pi(\overline{m-2}, \overline{n-1})$ will produce a sum of two squares of the form $2 \cdot m^{2l}$ for some $l \in \mathbb{N}$. What we have thus obtained is a restriction on possible witnesses which is independent of the particular cut-elimination strategy chosen. From this restriction one can see immediately that certain witnesses cannot be obtained from the above proof, for example odd sums of two squares or such that represent a Pythagorean triple. Also the more immediate argument that between n and $2 \cdot n$ (and thus below $m^2 \cdot n$) there is always a power of two (which is a sum of two squares) cannot be obtained for $m \geq 3$.

7 Conclusion

It is possible to extend the results of this paper to non-prenex sequents that contain only weak quantifiers. For that purpose, it suffices to replace the Herbrand-set by a suitable structure based on array formulas [4] or expansion trees [28]. An important direction for future work is to further tighten the constraints on the form of witness terms towards a characterization of the least upper bound. The structured terms introduced here provide an adequate technical basis for that purpose.

Another interesting prospect for further research consists in employing these grammars, or similar structures, for the analysis of mathematical proofs in the spirit of [25] and [26]. It would be very useful to obtain criteria on grammars that imply the existence of an Herbrand-disjunction fulfilling the growth conditions of [25] as those would guarantee that the Herbrand analysis provides a bound.

Acknowledgements I would like to M. Baaz, A. Carbone, P. Gerhardy, U. Kohlenbach, H. Schwichtenberg and an anonymous referee for important suggestions.

References

1. Matthias Baaz and Stefan Hetzl. On the non-confluence of cut-elimination. to appear.

2. Matthias Baaz, Stefan Hetzl, Alexander Leitsch, Clemens Richter, and Hendrik Spohr. Cut-Elimination: Experiments with CERES. In Franz Baader and Andrei Voronkov, editors, *Logic for Programming, Artificial Intelligence, and Reasoning (LPAR) 2004*, volume 3452 of *Lecture Notes in Computer Science*, pages 481–495. Springer, 2005.
3. Matthias Baaz, Stefan Hetzl, Alexander Leitsch, Clemens Richter, and Hendrik Spohr. CERES: An Analysis of Fürstenberg's Proof of the Infinity of Primes. *Theoretical Computer Science*, 403(2–3):160–175, 2008.
4. Matthias Baaz and Alexander Leitsch. On Skolemization and Proof Complexity. *Fundamenta Informaticae*, 20(4):353–379, 1994.
5. Matthias Baaz and Alexander Leitsch. Cut-elimination and Redundancy-elimination by Resolution. *Journal of Symbolic Computation*, 29(2):149–176, 2000.
6. Sam Buss. An Introduction to Proof Theory. In Sam Buss, editor, *Handbook of Proof Theory*, pages 2–78. Elsevier, 1998.
7. Alessandra Carbone. Turning cycles into spirals. *Annals of Pure and Applied Logic*, 96:57–73, 1999.
8. Alessandra Carbone. Cycling in Proofs and Feasibility. *Transactions of the American Mathematical Society*, 352:2049–2075, 2000.
9. H. Comon, M. Dauchet, R. Gilleron, C. Löding, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi. Tree Automata: Techniques and Applications. Available on: `http://www.grappa.univ-lille3.fr/tata`, 2007. release October, 12th 2007.
10. Pierre-Louis Curien and Hugo Herbelin. The Duality of Computation. In *Proceedings of the Fifth ACM SIGPLAN International Conference on Functional Programming (ICFP '00)*, pages 233–243. ACM, 2000.
11. Vincent Danos, Jean-Baptiste Joinet, and Harold Schellinx. A New Deconstructive Logic: Linear Logic. *Journal of Symbolic Logic*, 62(3):755–807, 1997.
12. Harold M. Edwards. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*. Springer, 1977.
13. Jean Gallier. Constructive Logics. Part I: A Tutorial on Proof Systems and Typed λ-Calculi. *Theoretical Computer Science*, 110(2):249–339, 1993.
14. Ferenc Gécseg and Magnus Steinby. Tree Languages. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages: Volume 3: Beyond Words*, pages 1–68. Springer, 1997.
15. Gerhard Gentzen. Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, 39:176–210,405–431, 1934–1935.
16. Gerhard Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, 112:493–565, 1936.
17. Philipp Gerhardy. Refined Complexity Analysis of Cut Elimination. In Matthias Baaz and Johann Makowsky, editors, *Computer Science Logic (CSL) 2003*, volume 2803 of *Lecture Notes in Computer Science*, pages 212–225. Springer, 2003.
18. Philipp Gerhardy. The Role of Quantifier Alternations in Cut Elimination. *Notre Dame Journal of Formal Logic*, 46(2), 2005.
19. Jean-Yves Girard. *Proof Theory and Logical Complexity*. Elsevier, 1987.
20. Kurt Gödel. Über eine noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958.
21. Willem Heijltjes. Proof Forests with Cut-Elimination Based on Herbrand's Theorem. In *Classical Logic and Computation (CL&C) 2008, participant's proceedings*. available at `http://wwwhomes.doc.ic.ac.uk/~svb/CLaC08/programme.html`.
22. Stefan Hetzl. Describing proofs by short tautologies. *Annals of Pure and Applied Logic*, 159(1–2):129–145, 2009.
23. David Hilbert and Paul Bernays. *Grundlagen der Mathematik II*. Springer, 1939.
24. Ulrich Kohlenbach. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer, 2008.
25. Georg Kreisel. Finiteness theorems in arithmetic: An application of Herbrand's theorem for Σ_2-formulas. In J. Stern, editor, *Logic Colloquium 1981*, pages 39–55. North-Holland, 1982.
26. Horst Luckhardt. Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken. *Journal of Symbolic Logic*, 54(1):234–263, 1989.

27. Richard McKinley. Herbrand expansion proofs and proof identity. In *Classical Logic and Computation (CL&C) 2008, participant's proceedings*. available at `http://wwwhomes.doc.ic.ac.uk/~svb/CLaC08/programme.html`.
28. Dale Miller. A Compact Representation of Proofs. *Studia Logica*, 46(4), 1987.
29. V.P. Orevkov. Lower bounds for increasing complexity of derivations after cut elimination. *Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta*, 88:137–161, 1979.
30. Dag Prawitz. *Natural Deduction: A Proof-Theoretical Study*. Almqvist and Wicksell, Stockholm, 1965.
31. Pavel Pudlák. The Lengths of Proofs. In Sam Buss, editor, *Handbook of Proof Theory*, pages 547–637. Elsevier, 1998.
32. Diana Ratiu and Trifon Trifonov. Exploring the computational content of the Infinite Pigeonhole Principle. to appear in the Journal of Logic and Computation.
33. Helmut Schwichtenberg. Proof Theory: Some Applications of Cut-Elimination. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 867–895. North-Holland, 1977.
34. Richard Statman. Lower bounds on Herbrand's theorem. *Proceedings of the American Mathematical Society*, 75:104–107, 1979.
35. W.W. Tait. Normal derivability in classical logic. In J. Barwise, editor, *The Syntax and Semantics of Infinitary Languages*, volume 72 of *Lecture Notes in Mathematics*, pages 204–236. Springer, 1968.
36. Gaisi Takeuti. *Proof Theory*. North-Holland, Amsterdam, 2nd edition, March 1987.
37. Christian Urban. *Classical Logic and Computation*. PhD thesis, University of Cambridge, October 2000.
38. Christian Urban and Gavin Bierman. Strong Normalization of Cut-Elimination in Classical Logic. *Fundamenta Informaticae*, 45:123–155, 2000.
39. Wenhui Zhang. Cut elimination and automatic proof procedures. *Theoretical Computer Science*, 91:265–284, 1991.
40. Wenhui Zhang. Depth of proofs, depth of cut-formulas and complexity of cut formulas. *Theoretical Computer Science*, 129:193–206, 1994.
41. J. Zucker. The Correspondence Between Cut-Elimination and Normalization. *Annals of Mathematical Logic*, 7:1–112, 1974.

A Appendix

A.1 Proof Reductions in First-Order Logic

The proof reduction rules for first-order logic consist of the those described in Section 2 for the quantifiers and for contraction and the following. For the case of the cut formula being introduced by weakening,

$$\dfrac{\dfrac{\overset{(\pi_1)}{\Gamma \to \Delta}}{\Gamma \to \Delta, A}\ \mathrm{w_r} \quad \overset{(\pi_2)}{A, \Pi \to \Lambda}}{\Gamma, \Pi \to \Delta, \Lambda}\ \mathrm{cut} \qquad \mapsto \qquad \dfrac{\overset{(\pi_1)}{\Gamma \to \Delta}}{\Gamma, \Pi \to \Delta, \Lambda}\ \mathrm{w}^*.$$

If the cut formula appears in an axiom,

$$\dfrac{A \to A \quad \overset{(\pi)}{A, \Gamma \to \Delta}}{A, \Gamma \to \Delta}\ \mathrm{cut} \qquad \mapsto \qquad \overset{(\pi)}{A, \Gamma \to \Delta}\ .$$

If the cut formula is introduced by propositional rules on both sides immediately above the cut, then

$$\dfrac{\dfrac{\overset{(\pi_1)}{\Gamma_1 \to \Delta_1, A} \quad \overset{(\pi_2)}{\Gamma_2 \to \Delta_2, B}}{\Gamma_1, \Gamma_2 \to \Delta_1, \Delta_2, A \wedge B}\ \wedge_{\mathrm{r}} \quad \dfrac{\overset{(\pi_3)}{A, B, \Pi \to \Lambda}}{A \wedge B, \Pi \to \Lambda}\ \wedge_{\mathrm{l}}}{\Gamma_1, \Gamma_2, \Pi \to \Delta_1, \Delta_2, \Lambda}\ \mathrm{cut}$$

$\mapsto$

$$\dfrac{\dfrac{(\pi_2)}{\Gamma_2 \to \Delta_2, B} \quad \dfrac{\dfrac{(\pi_1)}{\Gamma_1 \to \Delta_1, A} \quad \dfrac{(\pi_3)}{A, B, \Pi \to \Lambda}}{B, \Gamma_1, \Pi \to \Delta_1, \Lambda}\ \text{cut}}{\Gamma_1, \Gamma_2, \Pi \to \Delta_1, \Delta_2, \Lambda}\ \text{cut} \quad .$$

The other propositional connectives are treated analogously. We now turn to the rule permutations. For any unary rule ρ,

$$\dfrac{\dfrac{(\pi_1)}{\Gamma \to \Delta, A} \quad \dfrac{(\pi_2)}{\dfrac{A, \Pi' \to \Lambda'}{A, \Pi \to \Lambda}\ \rho}}{\Gamma, \Pi \to \Delta, \Lambda}\ \text{cut} \qquad \mapsto \qquad \dfrac{\dfrac{(\pi_1)}{\Gamma \to \Delta, A} \quad \dfrac{(\pi_2)}{A, \Pi' \to \Lambda'}}{\dfrac{\Gamma, \Pi' \to \Delta, \Lambda'}{\Gamma, \Pi \to \Delta, \Lambda}\ \rho}\ \text{cut}$$

which is a proof as regularity ensures that the eigenvariable condition cannot be violated. Similarly, for any binary rule ρ,

$$\dfrac{\dfrac{(\pi_1)}{\Gamma \to \Delta, A} \quad \dfrac{(\pi_2)}{A, \Pi_1 \to \Lambda_1} \quad \dfrac{(\pi_3)}{\Pi_2 \to \Lambda_2}\ \rho}{\Gamma, \Pi \to \Delta, \Lambda}\ \text{cut} \quad \mapsto \quad \dfrac{\dfrac{\dfrac{(\pi_1)}{\Gamma \to \Delta, A} \quad \dfrac{(\pi_2)}{A, \Pi_1 \to \Lambda_1}}{\Gamma, \Pi_1 \to \Delta, \Lambda_1}\ \text{cut} \quad \dfrac{(\pi_3)}{\Pi_2 \to \Lambda_2}}{\Gamma, \Pi \to \Delta, \Lambda}\ \rho .$$

All the obvious symmetric variants of the above reductions are also included in $\mapsto$.

A.2 Proof Reductions in Peano Arithmetic

The proof reduction rules for Peano arithmetic are those of first-order logic together with those described in Section 5 for eliminating the inductions as well as the following for shifting evaluations. If the auxiliary formula of the evaluation is main formula of a unary rule, then (e.g. for $\exists_{\mathrm{r}}$)

$$\dfrac{\dfrac{\dfrac{\Gamma \to \Delta, F(r,s)}{\Gamma \to \Delta, \exists x F(x,s)}\ \exists_{\mathrm{r}}}{\Gamma \to \Delta, \exists x F(x,t)}\ \mathrm{v}_{\mathrm{r}} \qquad \mapsto \qquad \dfrac{\dfrac{\Gamma \to \Delta, F(r,s)}{\Gamma \to \Delta, F(r,t)}\ \mathrm{v}_{\mathrm{r}}}{\Gamma \to \Delta, \exists x F(x,t)}\ \exists_{\mathrm{r}} .$$

If it is main formula of a binary rule except induction, then (e.g. for $\wedge_{\mathrm{r}}$)

$$\dfrac{\dfrac{\dfrac{\Gamma \to \Delta, A(s) \quad \Pi \to \Lambda, B(s)}{\Gamma, \Pi \to \Delta, \Lambda, A(s) \wedge B(s)}\ \wedge_{\mathrm{r}}}{\Gamma, \Pi \to \Delta, \Lambda, A(t) \wedge B(t)}\ \mathrm{v}_{\mathrm{r}} \qquad \mapsto \qquad \dfrac{\dfrac{\Gamma \to \Delta, A(s)}{\Gamma \to \Delta, A(t)}\ \mathrm{v}_{\mathrm{r}} \quad \dfrac{\Pi \to \Lambda, B(s)}{\Pi \to \Lambda, B(t)}\ \mathrm{v}_{\mathrm{r}}}{\Gamma, \Pi \to \Delta, \Lambda, A(t) \wedge B(t)}\ \wedge_{\mathrm{r}},$$

possibly dropping one of the two new evaluation rules if s does not appear in one of A, B. If it is main formula of a weakening, then

$$\dfrac{\dfrac{\dfrac{\Gamma \to \Delta}{\Gamma \to \Delta, A(s)}\ \mathrm{w}_{\mathrm{r}}}{\Gamma \to \Delta, A(t)}\ \mathrm{v}_{\mathrm{r}} \qquad \mapsto \qquad \dfrac{\Gamma \to \Delta}{\Gamma \to \Delta, A(t)}\ \mathrm{w}_{\mathrm{r}} .$$

If it is main formula of a contraction, then

$$\dfrac{\dfrac{\dfrac{\Gamma \to \Delta, A(s), A(s)}{\Gamma \to \Delta, A(s)}\ \mathrm{c}_{\mathrm{r}}}{\Gamma \to \Delta, A(t)}\ \mathrm{v}_{\mathrm{r}} \qquad \mapsto \qquad \dfrac{\dfrac{\dfrac{\Gamma \to \Delta, A(s), A(s)}{\Gamma \to \Delta, A(s), A(t)}\ \mathrm{v}_{\mathrm{r}}}{\Gamma \to \Delta, A(t), A(t)}\ \mathrm{v}_{\mathrm{r}}}{\Gamma \to \Delta, A(t)}\ \mathrm{c}_{\mathrm{r}} .$$

Furthermore, for any unary rule ρ,

$$\dfrac{\dfrac{\dfrac{\Gamma \to \Delta, A(s)}{\Gamma' \to \Delta', A(s)}\,\rho}{\Gamma' \to \Delta', A(t)}\,\mathrm{v_r}}{} \quad \mapsto \quad \dfrac{\dfrac{\dfrac{\Gamma \to \Delta, A(s)}{\Gamma \to \Delta, A(t)}\,\mathrm{v_r}}{\Gamma' \to \Delta', A(t)}\,\rho}{}$$

and similarly, for any binary rule ρ,

$$\dfrac{\dfrac{\dfrac{\Gamma_1 \to \Delta_1, A(s) \quad \Gamma_2 \to \Delta_2}{\Gamma \to \Delta, A(s)}\,\rho}{\Gamma \to \Delta, A(t)}\,\mathrm{v_r}}{} \quad \mapsto \quad \dfrac{\dfrac{\dfrac{\Gamma_1 \to \Delta_1, A(s)}{\Gamma_1 \to \Delta_1, A(t)}\,\mathrm{v_r} \quad \Gamma_2 \to \Delta_2}{\Gamma \to \Delta, A(t)}\,\rho}{}.$$

A normal form w.r.t. the rules for shifting evaluations is a proof where evaluations appear only below axioms and below inductions.